面向精准化服务的图书馆用户画像模型构建

■ 于兴尚[1,2] 王迎胜[2]

[1] 广州工商学院 佛山 528100 [2] 黑龙江大学信息资源管理研究中心 哈尔滨 150080

摘要：[目的/意义]基于用户画像构建图书馆精准服务模型，以期为用户进行精确服务提供参考，并为图书馆界进行精准化服务提供新方向。[方法/过程]以用户基本数据、用户动态行为数据、心理特征数据、情境感知数据、社交网络数据5个维度作为数据来源，基于用户画像技术构建数据来源层、数据处理层、用户数据仓库、精准应用层，打造在线服务推送、个性化推荐、用户流失预警、用户关系管理、场景化服务的服务特色。[结果/结论]虽然用户画像技术对于用户精准服务有重大的利用价值，但是用户隐私安全也应当考虑在内。所以从技术、用户意识与隐私政策、多维度的角度保护用户隐私，加上用户画像技术在图书馆的实施，通过这些措施达到图书馆精准服务的目标。

关键词：用户画像 服务模型 精准服务 隐私安全

分类号：G250.76

DOI:10.13266/j.issn.0252-3116.2019.22.005

1 引言

当前泛在信息背景下，信息混杂对图书馆服务能力和水平提出了更高的标准和要求，用户关注的重点已经从简单发现信息转变为以多元化方式取得满足自身需求的个性化信息资源，图书馆利用自身的资源优势并通过精准服务的方式为图书馆转型升级提供了可靠的保障，精准服务研究在图书情报界也引起了广泛的关注。从理论和模型构建来看，刘庆麟[1]以图书馆小数据概念、构成类型等相关理论基础作为出发点，提出以小数据为基础的读者实时兴趣模型。孙彤[2]从KBI(知识信息)的研究视角出发，结合国内外医学图书馆学科服务的差异，构建了基于知识信息的精准化学科服务模型。赵洪波[3]引入营销学中的RFM模型，构建图书馆精准服务平台。从研究对策上来看，牛勇[4]针对图书馆读者个性化服务流程中出现的资源分配有限以及图书馆界自身存在的社会效益评价问题，提出了一对一资源推荐服务对策。陈淑英等[5]针对"图书足迹"演化历程，并借助数学算法绘制三维空间结构图，进而揭示读者深层次的信息行为规律。从技术层面来看，马晓亭[6]、赵宁等[7]认为借助大数据技术进行数据分析既能改善图书馆服务，也是加强图书馆精准服务的有效措施。李剑[8]在谈到高校图书馆学科服务时，强调移动互联网以及云计算等相关技术与用户需求精准辨别深度融合，使得学科服务朝着精准化、系统化方向发展。综观图书馆精准服务覆盖内容，其研究主要侧重于结合其他学科理论基础以及图书馆职能定位、信息推荐、学科服务、发展对策等方面展开，实践研究成果并不凸显。其研究的重点没有从用户视域出发，没有将用户人口学信息、用户信息行为偏好、资源搜索方式、信息需求心理等多样环节考虑在内，并进行无差别推荐。图书馆如何根据用户阅读偏好、兴趣爱好、用户特征为用户提供个性化的精准服务，细化资源分配，用户画像不失为有效的解决方法之一。

当前用户画像是依据用户信息并对其抽象化，运用聚类、关联等数据挖掘方法进行信息可视化的处理过程[9]，从而把握不同用户的兴趣偏好，间接地根据用户需求变动进行及时反馈。目前"用户画像"在图书情报等相关领域有所涉及，如丁伟等[10]运用用户画像的基本原理分析用户数据，并通过数据样本实例验证用户画像在个人征信方面的应用。刘海等[11]以三枪集团的营销数据为对象，基于用户画像技术对消费群体进行细化，达到精准营销的目的。赵曙光[12]针对高转换率的社交媒体用户，通过深度

作者简介：于兴尚(ORCID:0000-0002-3763-3958)，硕士，E-mail:1170968754@qq.com；王迎胜(ORCID:0000-0001-6071-8776)，教授，硕士生导师。

收稿日期:2019-01-02 修回日期:2019-03-30 本文起止页码:41-48 本文责任编辑:王传清

访谈的方式,借助用户画像研究用户不同类型。刘海鸥等[13]借助用户画像技术的原理机制,以游客的基本信息、游客行为偏好信息、情景信息为基础,构建个性化的游客景点情景推荐模型。张钧[14]在人工智能的背景下,对构建的图书馆用户画像知识发现模型进行了详细分析。潘宇光[15]为了实现读者信息需求的精准性,针对高校智慧图书馆提出了用户画像系统读者信息需求模型。经检索发现,当前图书馆界关于用户画像的研究主要集中在图书馆应用、场景推荐、综述述评等领域,但是涉及到精准服务层面的研究少之又少。本文研究的主题是以用户画像为技术基础,构建图书馆精准服务模型,一方面可以针对各种各样的用户需求做出及时的反应,在用户相关信息分析与筛选中不断挖掘新的契机,扩展图书馆服务新空间;另一方面用户画像用于分析目标用户的原型,描述并勾勒用户行为、价值观以及用户需求,通过数据采集、行为攫取、模型构建等流程,更好地服务各类型用户,从而为图书馆精准营销提供额外助力和新思路。

2 用户画像及其作用

用户画像即用户信息标签化,通过收集、整理与分析用户的社会属性、专业背景、行为数据等主要信息和数据之后,以此为基准进行模型化构建,完美地抽象出用户的信息全貌。用户画像通过真实的数据将真实的用户与虚拟模型关联起来,对用户数据进行整理、清洗、组织、管理以及记录并反馈用户行为动态信息。通过数据收集为用户数据结构化开通了道路,用户画像凭借多维度的数据分类,连续叠加产生饱满的属性标签用来洞察用户的各种行为特点,然后利用数据挖掘、机器语言处理、聚类算法等技术对用户数据进行建模,最后以生动化、形象化的方式将用户信息以可视化的形态展现出来,通过用户画像建模可以抽取用户兴趣内容,打造生动、活泼的用户特征,明确地展示用户的潜在需求、行为偏好等信息动态,为图书馆员进行精准服务营销并识别目标用户开创了全新的方式,不仅缩短了用户获取信息的时间,还为图书馆服务升级提供了可视化的决策支持,从而在资源服务、用户粘性、服务重构等方面大幅度改善。

3 多维度采集数据构建用户画像

真实数据的多样化是建立用户画像的关键,依据多维数据组成的虚拟用户画像,其结果是不同用户在被细分之后的差异化标签描述[16],因此采集数据的容量决定了用户画像构建的清晰程度。用户数据主要来源于图书馆门户网站、图书检索平台、移动阅读终端、门禁系统数据库、图书馆综合管理系统、社交网络APP、可穿戴设备等,主要采集用户基本数据、用户动态行为数据、社交网络数据、心理特征数据、情境感知数据五大类型。

在图书馆门户网站中存储了大量的用户基本数据,主要包括姓名、年龄段、学号、籍贯、民族、专业、教育层次、专长喜好、婚姻状况、注册信息等;用户动态行为数据又可分为显性可变数据和隐性可变数据,其中图书检索平台和移动阅读终端包含了显性可变数据(活跃频次、主题分享、订阅渠道、访问时间、行为偏好、下载频次、最新评价、打分、登录频率、检索途径等)和隐性可变数据(页面浏览、内容收藏、停留时间、访问程度、检索速度等);除了上述数据以外,门禁系统数据库中记录的用户刷卡信息(刷卡地点、频次、时间等)以及图书馆综合管理系统中借阅次数、图书归还状态、活动预约等也可归于用户显性数据。社交网络APP(微信、QQ、微博等)可以搜集到用户社交网络数据(互动分享、业务咨询、微博动态、邮件发送等),同时管理信息系统和可穿戴设备可搜集用户阅读和检索过程中的心理特征(社交表情、省力心理、查准心理、访问偏好、文献喜好倾向等);最后情境感知数据(馆内温度、湿度、光感强弱、周围环境、LBS位置)可通过传感设备(温度传感器、湿度传感器、位置传感器、RFID)等获取。由于以上5种数据涉及多种系统平台,各系统之间关联程度和统一程度不紧密,这就需要考虑到数据融合的问题,这也是精准服务构建需要解决的主要问题之一。目前,通用的解决方法是通过用户智能一卡通账号作为登录端口,这样使得整合图书馆各类系统和图书馆资源中的数据变得不再复杂,各类系统上用户相关的数据痕迹一起构成了用户画像构建的多样数据来源。

鉴于图书馆用户数据具有小数据、更新快、维度高、形式多样等特点,依据相关性准则、适度性原则并不需要分析用户的全部数据,这样既规避了多余数据的干扰,又提高了用户画像应用的价值密度。用户画像基于空间全方位、时间全过程、深度精准化的特性展示和预测用户的行为模式,从而为图书馆决策提供高质量的服务。以上数据类型共同组成了用户画像建模的基础,如图1所示:

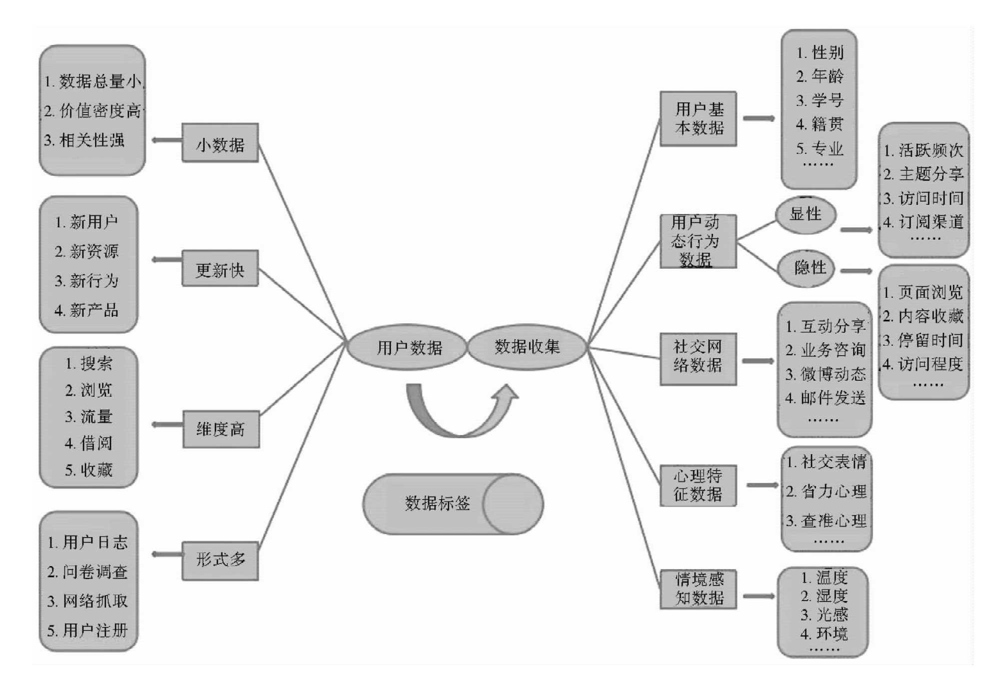


图1　图书馆用户画像多维数据源

4　用户画像精准服务模型的规划设计及搭建

4.1　模型构建的规划设计

笔者基于人体构造的相关理论，将生命呼吸、骨骼、神经传导系统、血液4个元素作为参照对图书馆服务层次结构进行重新规划和设计。信息技术背景下基于用户画像的图书馆精准服务同样就像一个完整的生命体，用户信息流如同生命气息供应模型搭建的基础来源；信息的组织和处理如人体骨骼，支撑服务系统的整体布局；信息的存储和分析如神经传导系统，负责信息元的有效传输；图书馆服务如血液，负责服务流动的整体方向。基于此构建图书馆用户画像精准服务模型，包括数据来源层、数据处理层、用户数据仓储和精准应用层，分别对应人体结构中的生命呼吸、人体骨骼、神经传导系统和人体血液。其中，数据来源层负责提供精准服务的信息源；数据处理层功能在于提供数据处理的关键技术和组织用户数据；用户数据仓储主要负责数据处理层中传递信息的存储和分析；精准应用层是四大层次中的重点层次，也是图书馆精准化服务的灵魂，为用户提供决策化、智慧化的信息服务。层次结构的多样性紧密相连，保证该体系在变化莫测的数据环境中具有较好的可塑性和延展性。

4.2　用户画像精准服务模型的搭建

图书馆精准服务离不开技术层次的构建，如何在数据驱动环境下实现图书馆各个要素间的优化与整合，构建层次化的精准服务体系是当前图书馆界研究的核心要点。在此体系的推动下，用户在信息需求行为的驱动下访问图书馆，图书馆中的各种服务管理系统将感知、识别、整合用户浏览图书馆过程中产生的数据，并上传数据处理层经自然语言处理技术、数据挖掘、社会网络分析、感知技术、音视频内容处理技术、知识图谱等进行过滤、清洗和组织处理。对处理的数据进行加工存储，通过 Hive、MR、Spark 离线计算技术进行用户特征抽取建立特征工程，借助 Storm、Storm steaming 实时计算技术构建用户模型，同时根据用户稳定数据、用户可变数据、社交网络数据、心理特征数据、情境感知数据五大类型数据进行特征分析，运用分类、聚类、回归、预测等方法深度挖掘用户对相似信息的利用、收藏等行为，计算不同用户之间的相似程度并以可视化的形式突出易懂、易用的个体以及用户群体之间的用户画像；最后预测用户信息需求在顶层交互端口开展个性化的信息推荐服务，也基于用户基本信息、社交数据进行相应的在线服务，建立特殊的预警机制，区分核心用户、普通用户和潜在用户；为防止用户流失、规范用户服务机制、制定相应的服务对策奠定基础保障。用户心理数据和情景化数据是进行用户关系管理和场景化推荐的基础，可以极大改善图书馆服务绩效，有利于分析用户隐藏需求，挖掘潜在用户，有助于改进图书馆网页设计。另外，用户画像构建不是固定不变的，用户需求以及兴趣偏好等用户信息会随着用户阅历的增加而改变，及时的需求反馈和更新是图书馆进

行服务营销的有效手段。笔者基于用户画像构建的技术层次,从数据来源层、数据处理层、用户数据仓库、精准应用层 4 个层面对用户画像的技术模型进行组建(见图 2),从关于用户人口结构的数据、行为、社交、情境、情感的数据中演化关联数据并依此甄别用户的行为倾向以及属于隐藏领域的某些线索。

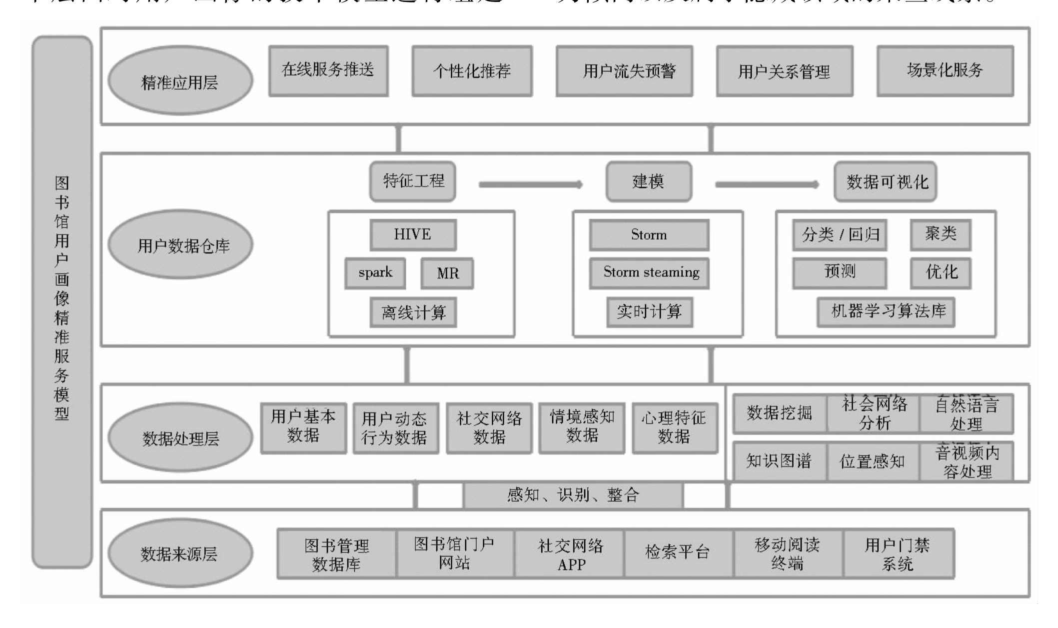


图 2　图书馆用户画像精准服务模型

4.2.1　数据来源层　数据来源层是精准服务的基础设施层,由图书管理数据库、图书馆门户网站、社交网络 APP、检索平台、移动阅读终端以及用户门禁系统等组成,主要提供用户画像分析的数据源并完成用户人口数据、用户阅读情境、用户服务数据反馈等数据的收集,通过网络连接将用户、服务设施、管理系统互联成统一的整体,以实现图书馆服务环境、用户、设施等要素之间产生的数据的感知、识别以及整合。

考虑到用户数据采集的分散性以及各大图书馆管理系统获取数据的独立性,因此需要整合不同系统间的数据,再进行用户画像的描绘[17]。为了保证数据的时效性和管理的统一性,可将用户智能一卡通作为数据串联的接口,将图书馆用户管理平台、腾讯 QQ、检索平台与 RFID 门禁网络进行功能兼容,达到整合用户基础信息和社交网络数据的效果。由数字图书馆、学科信息门户、搜索引擎、图书检索系统统一用户会话数据、用户信息行为特征等,情景数据可利用传感器技术、智能手机、位置感知技术等整体采集图书馆情景状态信息,确保用户信息采集的可靠性、完整性以提高用户数据分析的信用度。

4.2.2　数据处理层　数据处理层是由用户数据和相关的处理技术两部分构成,主要功能是对用户数据进行组织与挖掘。图书馆用户数据是图书馆日常服务流程与管理过程中产生的数据流,按照用户数据分类大致可分为以下 5 类:①用户基本数据,包括用户人口统计信息(性别、年龄、专业、教育背景等);②用户动态行为数据,包括用户借阅频次以及浏览、收藏、检索过程中产生的图片、文本资源、视频、音频等数据集合;③社交网络数据,包括互动数据、转发动态、网页访问次数等;④情境感知数据,包括用户距离、位置状态、光感强弱、环境等数据;⑤用户心理特征数据,包括信息需求意愿强度、检索工具使用省力心理、服务环境满意度以及资源倾向等。由于图书馆用户数据涉及文本信息、图片、音频、视频、用户心理特征等复杂变化的数据态势,可使用自然语言处理技术、知识图谱研究用户基本数据、行为可变数据,凭借音视频内容处理技术分析图片、音频、视频等非结构化数据结构[18],通过数据挖掘、社会网络分析技能处理用户社交网络数据和用户心理趋势[19],使用感知技术探索情景感知数据和馆室气候,挖掘用户位置信息进行情境化服务,根据气候变化进行相关的温度、光度调整。总之,图书馆采用灵活多样的新兴技术对图书馆用户数据的管理和实时控制,为后续服务创造技术条件。

4.2.3　用户数据仓库　数据仓库是将分散、结构不统一的数据转换成同构数据格式,为管理者进行决策分析提供数据存储环境[20],对图书馆决策者进行创新服务并判断其效益具有重大的参考价值。因此,图书馆设立用户数据仓库使用户信息处理结果具有易处理、

于兴尚，王迎胜.面向精准化服务的图书馆用户画像模型构建[J].图书情报工作,2019,63(22):41－48.

广泛使用性等特性，以满足图书馆服务评估的各种需求[21]。Hive作为一种数据仓库技术，用来存储和处理用户在线点击、评论、浏览等大量阅读行为信息，也是进行特征工程构建的关键。通过对用户信息行为进行精确分析可以统计用户的行为特征和计算阅读时间，借助离线计算功能，从而准确得出用户真实的阅读行为，也为用户在线服务推荐提供了保障。MR（混合现实技术）的引入为虚拟、现实和用户之间搭建了一座交互反馈的信息桥梁，以增加用户服务体验的真实感。加上用户对实时性查询要求越来越高，Spark与Hive的有效结合为提高实时查询速度带来了可能，杨宁、黄婷婷通过研究发现Hive on Spark的运行模式下执行的检索速度与Hadoop模式相比提高了17.42－46.35倍[22]。Storm和Spark steaming都是高通量、容错处理的分布流式处理系统，用于用户信息流进行批量处理、实时计算，监控信息节点存储位置，并进行信息标记、跟踪信息流向，纠正错误信息状态，建立良好的容错机制。Storm和Spark steaming的容错功能为信息传输和存储的准确性提供了更好的支持，两者的有效使用可以查缺补漏，增强信息处理的可信性，对用户建模、提高个性化推荐服务的准确性意义重大。为了增加图书馆用户画像与服务系统的匹配幅度，充分使用自然语言信息以及分类、回归、聚类、预测、优化等机器学习算法的引入对于解析和评估个体、群体用户画像、话题聚合、主题分类等用户画像分析层面效果显著，生动地诠释了用户画像的语义化、短本化等特点。

4.2.4　精准应用层　用户画像结合用户真实数据，增添描述性的信息元素、场景化信息、典型化特征信息到用户画像中，让用户画像更加真实、丰满、容易记忆，从而为图书馆在线服务推送、个性化推荐、用户流失预警、用户关系管理、场景化服务提供策略依据。

（1）在线服务推送。在用户检索信息的过程中，针对用户检索记录以及相关内容等线索，通过用户数据仓储匹配用户信息需求，进行在线服务推送，这一功能不仅可以从内容上提供服务，还可以关联用户的信息行为（学科领域、检索历史等）进行操作，并将处理后的相关信息在线推荐给用户。通过对用户的信息需求分析，利用用户画像技术，从用户的视角实现无序信息流向有序信息流的转换，针对不同用户提供不同的在线推荐服务，如图书力荐、新书推送、学科知识服务等多种在线推荐服务，以高质量知识形态呈现在读者面前。

（2）个性化推荐。应用用户画像技术以用户个性化需求为推荐准则，所涉及的功能即个性化推荐。基于用户在图书馆虚拟环境中产生的借阅数据、检索下载、账号活跃度、评论等数据，运用用户画像技术精确描述用户行为全貌以及成功预测用户信息需求，为用户提供个性化信息订制，让用户的资源体验从技术的表面交流上升到情感层面的互动。来自德国卡尔斯鲁厄大学的A. Geyer-Schulz、A. Neumann和A. Thede[23]通过测量用户使用同样文献共出现的频次，以用户行为为基准构建了用户行为画像信息推荐系统，并将该技术成功应用于学校OPAC推荐系统服务中。可见，用户画像技术与图书馆个性化服务的相互融合，在已知的信息资源中分析读者行为兴趣，个性化地为用户推送信息，不仅弥补了推荐结果的规范性和通用性缺陷，还有利于跨学科、跨领域不同资源之间协同发展。

（3）用户流失预警。用户流失预警可以帮助图书馆明确哪些用户流失、用户等级划分以及流失的原因，并为以后图书馆优化资源与服务提供借鉴。图书馆的所有用户都具有生命周期的特性，用户的生命周期是指用户从开始注册成为图书馆使用者到不再利用图书馆的整个过程[24]。通过用户画像技术分析用户相关数据，如用户搜索信息的方式、服务依赖程度、用户最近使用图书馆服务的时间、服务类型以及访问次数等，判断用户所处的生命周期，并聚类分析用户的行为属性以及用户流失的原因，根据不同的流失原因设置针对性的对策，确定应该挽回哪类用户，具体采取什么措施，将普通用户转换为核心用户和忠诚用户，达到精准服务营销的效果。

（4）用户关系管理。用户关系管理就是了解用户心理需求，分析和探索用户信息行为，以便更好地服务用户[25]。熟知用户心理可以发现用户的需求倾向，研究用户的行为内在变化以及发展规律，从而减少用户服务的被动性。借助用户画像技术收集用户相关心理数据，通过聚类、关联用户求准、求全、求快、检索便捷、需求层次性、信息偏好、用户满意度等心理，挖掘用户记录信息以及会话交互数据，提炼用户心理诉求规律，可以针对不同层次的需求来制定针对性的服务措施，提升用户自我效能和用户信息素养。

（5）场景化服务。场景化推荐模式基于场景化数据，并对目标用户进行画像描述。精准服务模型通过用户基础数据库、智能设备、网络状态等内外环境获取数据，然后精准服务层定位获取、无缝衔接，利用推荐服务引擎为用户推荐知识单元。如2017年12月20日下午2点55分，某管理科学与工程专业的研究生经

过图书馆借阅室二时,操作 iPhone 6S 终端登录图书检索系统查找《运筹学基础》一书,图书馆利用用户画像技术收集的相关情景信息,根据读者位置和搜索时间精准将相关信息推荐给该用户。

5 精准化服务过程中用户隐私保护对策

在当今时代,由于用户进入互联网、社交网络、物联网的足迹愈加深入,图书馆对用户信息了如指掌,甚至达到了资源化的地步。加上信息收集、多元异构数据融合等多种技术元素的不断渗透,用户隐性数据的泄露往往不可知、难以控制[26]。所以如何保证用户信息的安全性,不仅仅是技术的问题,更应该兼容用户隐私安全意识与政策协调发展,以多维度服务思维模式关注用户隐私。

5.1 网络环境下用户隐私技术保护对策

用户通过网络获取图书馆资源或者享用图书馆各种服务时,其注册、登录虚拟图书馆的基本信息、背景信息、行为记录易于在网络环境下出现被窃取、破坏等风险问题,图书馆可采用 EPAL 语言和 P3P 协议标准等保护机制来设置信息搜索的类型、目的等,以避免这些问题的发生[27]。在数据处理层面上,为了达到服务效果并隐藏用户行为踪迹,可采用用户匿名化的长效机制,使得用户信息很难定位,可以达到用户信息处理和隐私保护的双重效果。为了防止非法侵入用户数据仓库等黑客行为,可利用加密技术保护用户信息明文,加密后的信息让非法侵入者的攻击毫无作用。针对精准服务层存在的身份认证和访问控制机制严谨失控等问题,可通过密码静态和动态双因素认证来保障用户身份唯一性。除此之外,访问权限控制手段的实施在监控用户的合法性、入网时间、登录用户人数等方面效果显著,因为可严格控制合法用户的非权限操作,进而保护用户信息安全。

5.2 用户主观感知的有效性与隐私政策保护客观性相融合

图书馆在为用户提供资源和服务的同时,应该告知用户信息采集、处理、分析等使用的目的和用途,在征得用户同意之前,不应该擅自使用,以消除用户存在的后顾之忧。图书馆应大力宣传信息安全意识,加强用户自身主观感知信息的敏感度,认识到保护信息的有效性和重要性。在服务流程中完善、健全的隐私政策能给予用户信息全面的保护,在图书馆明显位置公布个人隐私信息的保护公告,增加用户信息使用的公开透明度。图书馆应该实施明确的隐私信息保护制度,加强对图书馆员的素质教育,深化馆员对用户信息安全保护的职业操守。构建"图书馆 - 图书馆员 - 服务终端提供商"三方联合机制,并签订用户隐私安全保护协议,一旦发生信息泄露事件,对相关涉事人员或者单位进行经济处罚并追究其相应的法律责任。

5.3 以多维度的视角探索用户隐私保护

用户隐私保护是学术界关注的热点话题,仅从技术和用户隐私政策的范围来保障和维护用户隐私往往是不够的,多维度、动态化的隐私协商措施,对于用户隐私保护政策设计的灵活性、可选择性建立必要保障。图书馆应利用病毒查杀、网络漏洞修复、数据备份、选择拒绝或者接受 Cookies 的方式来防范由于计算机硬件软件损害、自然灾害、人为因素等造成的不可抗拒后果,对于完善安全预测与分析、建立深层的安全防御体系都具有重大意义。此外,开展多学科的隐私信息融合举措,发挥学科优势,科学地剖析和解决用户隐私保护政策的缺陷问题也是完善用户隐私安全保护的有效手段。

6 结语

当前用户需求日渐多元化,图书馆用一种通用模式来服务用户,其产生的效用过于单一。在不能准确、全面分析和洞察用户的相关信息的前提下,多数个性化、适度化的服务策略缺乏事实依据,从而引发设计方案适用性和灵活性欠缺问题,难以有的放矢,用户的参与度不高,难以达到理想的效果。大数据和网络技术时代,图书馆可以利用大数据画像对用户的信息行为进行有效采集,高效率地挖掘、分析和运用用户的稳定信息以及用户在网络环境下产生的动态信息,通过数据过滤,清洗冗杂数据并对数据进行结构化组织,可视化呈现,并赋予适合用户的个性化标签,可以真实地反应用户不同时间的信息需求、阅读偏好和行为习惯,大大减少人为判断的主观性。由此绘制的用户画像可以较为透彻地描述用户的信息全貌。这必然会为图书馆服务的改革提供新的设计模式。与此同时,图书馆要想真正实现"以用户为中心",需尊重用户基本权益,舒缓用户对隐私数据的心理压力,制定适合用户的隐私保护手段,多方面帮助图书馆服务工作从一体化、整体化向个性化服务转变,使图书馆服务从宏观视野转向微观视野,最终实现数据转换价值的目的,精准达到"私人订制"的服务高度,实现服务内容的全覆盖。

本研究为图书馆精准服务提供了新的思路,也为图书馆建立品牌服务战略带来了一些参考价值,但是

研究过程中的一些问题也不容忽视:①用户参与的粘性问题。用户进入系统的参与频次是影响数据采集的重要因素。如果一些用户在一定时间内登录用户频次不正常、关注数量不足、用户提问不积极、个人信息全面性不高等,会造成用户服务目的性不强以及数据不真实等现象,因此用户数据达到绝对数量对于用户画像研究具有基础价值。②研究方法的综合性问题。本文主要从技术角度采集数据,对于数据完整性存在一定的欠缺,易于出现数据多样化不足和数据量偏小等问题,特别是用户心理特征数据的准确性难以把握。但是问卷调查和技术的有效统一,对于收集数据的合理性和保障服务的准确性具有不容忽视的价值。

参考文献:

[1] 刘庆麟. 基于小数据的图书馆精准服务研究[J]. 图书馆工作与研究,2017(5):45 –50.

[2] 孙彤. 以知识的信息(KBI)为视角论医院图书馆精准化学科服务[J]. 图书馆工作与研究,2016(9):59 –62.

[3] 赵洪波. 基于 RFM 模型的高校图书馆精准服务研究[J]. 情报探索,2016(12):26 –32.

[4] 牛勇. 图书馆精准服务研究[J]. 图书馆学研究,2016(5):50 –56.

[5] 陈淑英,张艳花,山洁,等. 基于“图书足迹”的高校图书馆精准服务模型构建及应用研究[J]. 图书馆杂志,2017(1):71 –74.

[6] 马晓亭. 大数据时代图书馆个性化智慧服务 QOS 保障研究[J]. 图书馆理论与实践,2016(2):38 –41.

[7] 赵宁,张玲,唐振宇. 移动互联网时代下高校图书馆学科服务[J]. 高校图书馆工作,2016 (2):67 –70.

[8] 李剑. “互联网 +”环境下高校图书馆精准化学科服务的构建[J]. 图书馆研究与工作,2017(3):83 –85.

[9] 胡媛,毛宁. 基于用户画像的数字图书馆知识社区用户模型构建[J]. 图书馆理论与实践,2017(4):55 –57.

[10] 丁伟,王题,刘新海,等. 基于大数据技术的手机用户画像与征信研究[J]. 邮电设计技术,2016(3):45 –96.

[11] 刘海,卢慧,阮金花,等. 基于“用户画像”挖掘的精准营销细分模型研究[J]. 丝绸,2015(12):1 –6.

[12] 赵曙光. 高转化率的社交媒体用户画像:基于 500 用户的深访研究[J]. 现代传播:中国传媒大学学报,2014, 36(6):115 –120.

[13] 刘海鸥,孙晶晶,苏妍妪,等. 基于用户画像的旅游情境化推荐服务研究[J]. 情报理论与实践,2018(5):58 –61.

[14] 张钧. 基于用户画像的图书馆知识发现服务研究[J]. 图书与情报,2017(6):56 –65.

[15] 潘宇光. 校智慧图书馆读者信息需求的用户画像[J]. 合肥工业大学学报(社会科学版),2018 (2) : 113 –119.

[16] 陈添源. 高校移动图书馆用户画像构建实证[J]. 图书情报工作,2018,62(7):85 –90.

[17] 刘海鸥,黄文娜,谢姝琳. 面向情境化推荐服务的图书馆用户画像研究[J]. 图书馆学研究,2018(20):58 –62.

[18] GANDOMI A, HAIDER M. Beyond the hype: big data concepts, methods, and analytics [J]. International journal of information management, 2015, 35 (2):137 –144.

[19] ZOU H, CHEN H M, DEY S. Understanding library user engagement strategies through Large -Scale Twitter analysis [C]//IEEE first international conference on big data computing service and applications. Washingten, DC:IEEE computer society, 2015:361 –370.

[20] 王伯恭. 中国百科大词典[M]. 北京:中国大百科全书出版社,1999:20.

[21] LAITINEN M, SAARTI J. A model for a library-management toolbox: data warehousing as a tool for filtering and analyzing statistical information from multiple sources[J]. Library management, 2012 (4/5):253 –260.

[22] 杨宁,黄婷婷. 基于 Spark 的大数据分析工具 Hive 的研究[J]. 计算机时代,2018(11):1 –8.

[23] GEYER-SCHULZ A, NEUMANN A, THEDE A. An architecture for behavior-based library recommender systems[J]. Information technology and libraries,2003,22(4):165 –174.

[24] 李政. 基于用户生命周期的图书馆用户管理研究[J]. 图书馆建设,2008(11):18 –23.

[25] 黄晓斌,朱俊卿. 数字图书馆用户的心理研究[J]. 图书馆学研究,2006(1):185 –193.

[26] 孟小峰,慈祥. 大数据管理:概念、技术与挑战[J]. 计算机研究与发展,2013,50(1) :146 –169.

[27] 潘浩,张幸. 一种基于自主计算的数字图书馆个性化服务隐私保护框架[J]. 图书情报工作,2009,53(11):75 –77.

作者贡献说明:

于兴尚:提出研究命题、研究思路,撰写与修改论文;
王迎胜:负责论文的整体布局和结构规划,协助修改论文。

Research on the Construction of Library Precision Service Model Based on User Portrait

Yu Xingshang[1,2] Wang Yingsheng[2]

[1] Guangzhou College of Technology and Business, Fushan 528100

[2] Information Resource Management, Research Center School of Information Management, Heilongjiang University, Harbin 150080

Abstract: [**Purpose/significance**] Constructing a precise service model based on user portraits, provides library a reference to perform precise services for users, and provide a new direction for the library industry to carry out precision

services. [**Method/process**] Based on the five dimensions of user basic data, user dynamic behavior data, psychological feature data, context-aware data and social network data, service features of online service push, personalized recommendation, user churn warning, user relationship management and scene service have been created based on data source layer, data processing layer, user data warehouse and precision application layer. [**Result/conclusion**] Although user portrait technology has significant value for users' accurate services, user privacy should also be considered. Therefore, from the perspective of technology, user awareness and privacy policy, multi-dimensional protection of user privacy, and the implementation of user portrait technology in the library, through these measures to achieve the goal of library precision services.

Keywords: user portrait service model precision service privacy security

"2019 科学计量与科技评价天府国际论坛"在成都举办

2019 年 11 月 11 - 12 日,由中国科学院成都文献情报中心主办、科学计量与科技评价研究中心(SERC)承办、9 家单位协办、12 家期刊支持的"2019 科学计量与科技评价天府国际论坛"(以下简称"天府论坛")在成都成功举办。论坛吸引了来自比利时、挪威、瑞典、英国、法国、爱尔兰等国及国内 60 余家单位的近 260 名相关领域专家学者参会,并通过特邀大会报告、大会主题报告、热点对话、投稿论文报告和专题学术沙龙等丰富的论坛形式,深入研讨和交流科学计量与科技评价领域的理论方法研究与实践应用等方面的最新进展及发展趋势。

在论坛开幕式上,中国科学院成都分院分党组书记王嘉图、中国科学院成都文献情报中心主任张志强、中国政法大学副校长时建中分别代表主管单位、主办单位、协办单位出席开幕式并致辞。开幕式由中国科学院成都文献情报中心科技处处长、SERC 执行主任陈云伟研究员主持,闭幕式由中国科学院成都文献情报中心副主任杨志萍研究员主持。

在大会特邀报告中,比利时鲁汶大学 Ronald Rousseau 教授作题为《Benchmarking & Ranking》的报告,首先对排名和基准的概念进行了界定,介绍了基准在科研及科学计量中的应用,并对全球具有典型性的科研排名方法进行了比较,指出期刊或科研机构等排名是促进机构发展以及支撑科技决策的必要手段,但在基准选择、排名方法制定和方法稳定性等方面仍然需要不断完善。英国伦敦国王学院 Grant Lewison 教授作题为《The evaluation of prostate cancer research through references on clinical practice guidelines》的报告,讨论了各国家在临床医学参考文献使用和数据库检索方面存在的显著差异问题。挪威北欧创新、研究和教育研究所 GunnarSivertsen 教授和 Dag Aksnes 教授分别围绕科学计量与科技评价的社会影响以及引文计量与科研质量评价,作题为《Understandingand Evaluation the Societal Impacts of Research》和《Citations, Citation Indicators, and Research Quality》的报告。瑞典乌普萨拉大学 PerAhlgren 副教授作题为《Enhancing Direct Citations: A Comparisonof Relatedness Measures for Community Detection in a Large Set of Pub Med Publications》的报告,提出一种利用直接引用关系计算大型学术文献集中的团体探测方法。法国斯特拉斯堡大学 Jean-Charles Lamirel 教授作题为《Time-based Science Analysis》的报告,以中国科学学四十年的数据为基础,提出了一种完全无监督的方法来分析大量异构文本数据的演化规律。爱尔兰国立都柏林大学 Lai Ma 助理教授作题为《The Challenges of Ex-Ante Impact Assessment: A Case from Science Foundation Ireland(SFI)》,从事前角度量化项目的影响力,指出项目影响力包括经济、社会、政策引导等多方面),评价应以同行评议为主开展多指标综合评价。武汉大学张琳教授作题为《Understanding and evaluating research and scholarly publishing in the social sciences and humanities》的报告,通过对国际数据库收录人文社科成果情况进行分析,对人文社科评价提出思考,指出国际现有主要文献数据库可能无法反应人文社科成果真正的学术价值,我国的人文社科评价还需进一步本土化。

中国科学院成都文献情报中心主任张志强研究员作题为《国际重大科学奖励的计量与评价》的特邀报告,从国际重大科学奖励的研究与发展、案例与分析、政策与建议等三个方面,论述了国际重大科技奖励制度的起源与发展,详细介绍了对诺贝尔科学奖以及主要科技领域国际重大奖励的计量分析结果,以及从国际重大科技奖励计量和评价研究中总结出的一系列科学规律,并提出了对我国科技奖励制度改革的重要政策建议,以及促进我国重大科学成果产出的相关科技政策建议等。

北京科学学研究中心张士运研究员作题为《京沪深三地科创中心评价》的特邀报告,通过详细的指标量化评价和分析,指出北京是以高校所主导的技术驱动型发展模式,上海是以传统产业和高技术产业并存的混合经济发展模式,深圳是高技术产业为特征的发展模式。在热点对话环节,围绕"科技评价研究责任与担当"、"未来十年的科学计量学"以及"科学学大数据分析机遇与挑战"等三个热点对话主题,三组十多位科学计量与科技评价领域的优秀中青年学者围绕对话主题,分别充分阐释自己的学术观点和发展见解,并与参会代表展开了热烈的交流和讨论。

在投稿论文报告环节,有 28 位投稿论文作者分享了自己的论文成果,论坛组织现场专家评审出了优秀论文。

天府论坛还举办了由南京理工大学、大连理工大学和北京理工大学研究团队共同发起的"全文本计量分析"专题学术沙龙。

论坛上还发布了由中国科学院成都文献情报中心研究团队撰写的《我国省域基础研究发展现状分析》、《DNA 存储科技国际发展趋势分析》及《中欧班列沿线国家知识产权政策风险研究》等最新研究报告。

中国科学院成都文献情报中心《科技咨询频道》
2019 年 11 月